PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04M

A2

(11) International Publication Number: WO 98/23075

(43) International Publication Date: 28 May 1998 (28.05.98)

(21) International Application Number: PCT/US97/21279

(22) International Filing Date: 20 November 1997 (20.11.97)

(30) Priority Data:
08/755,432 22 November 1996 (22.11.96) US

(71) Applicant: UNISYS CORPORATION [US/US]; Township Line and Union Meeting Roads, M/S C1SW19, Blue Bell, PA 19424-0001 (US).

(72) Inventors: CHAMPA, John, J.; 7800 Hartwell Street, Dearborn, MI 48126 (US). WILSON, George, S.; 40 W. Green Hill Road, Broomall, PA 19008 (US).

(74) Agent: STARR, Mark, T.; Unisys Corporation, Township Line and Union Meeting Roads, M/S C1SW19, Blue Bell, PA 19424-0001 (US).

(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MULTIMEDIA TELECONFERENCING BRIDGE

(57) Abstract

A hub (4) for multimedia multipoint video teleconferencing includes having a plurality of input/output ports (22), each of which may be coupled to a communication channel (12) for interchanging teleconferencing signals with remote sites (2). The hub (2) includes a plurality of signal processing functions (30) that can be selectively applied to teleconferencing signals so that the signals distributed by the hub (4) are in a form desired by the recipients or required by their equipment. Signal processing may include video, data, graphics, and communication protocol or format conversion, and language translation. The hub (4) may generate data (82) relating to the use of its processing capabilities during a teleconference, so that accounting or billing for a teleconference may be based at least in part on which hub resources were used, the extent of their use, and the person desiring their use. The identification of a signal processing function (30) to be used during a teleconference may be automatically performed in response to the content of signals received at the hub (4) during the teleconference.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MULTIMEDIA TELECONFERENCING BRIDGE

BACKGROUND THE OF INVENTION

This invention relates to teleconferencing systems. More particularly, this invention relates to a bridge system which enables a teleconference to occur between participants having a variety of purposes and requirements for the conference, at sites having a wide variety of equipment and communications facilities to conduct the conference.

Teleconferencing - holding a conference by telecommunicating among people in different locations - is becoming an increasingly important business tool. Teleconferencing reduces business travel cost and time lost to business travel, and eases difficulties in coordinating the schedules of those who are to meet. Video conferencing, or multimedia teleconferencing including video, has required specialized facilities to provide and display the video signals and to establish video communication channels between the meeting sites where the conference participants are located. For multipoint multimedia teleconferencing, equipment at several sites, which typically are dedicated video conference rooms, is coupled to a "hub" or "bridge". The hub receives signals from each site, selects one of the signals (such as the signal from a site where a participant is speaking), and distributes the selected signal to the other sites. Many business locations lack such specialized facilities and access to communication channels having sufficient bandwidth to support video teleconferencing, and so video teleconferencing has generally required people to travel to dedicated videoconference facilities having the necessary video equipment and communication access.

The opportunity for videoconferencing is increasing, in part due to increasing availability of suitable communication channels and increasing capability of image processing systems to transmit acceptable quality video in a given bandwidth. The proliferation of client/server computer systems having high speed personal computers with high resolution monitors on workers' desks, coupled to each other and to a server via a local area network, has made it possible to add video input and processing to provide "desktop videoconferencing." An example is the system sold by Intel under

the mark PROSHARE. While these systems have many important benefits that flow from the use of an already-installed information infrastructure to support video teleconferencing, that infrastructure also has many attributes that result in drawbacks for video teleconferencing. The high bandwidth required for even low quality video can consume an inordinate amount of network resources and substantially interfere with the functions for which the infrastructure was provided in the first place. ISDN telephone service has been required to enable a desktop video teleconferencing system to confer with an off-network site, and multipoint conferences require a hub to effect switching.

Use of an existing information infrastructure like a local area network also has an advantage in that the information system administrators can control access to and prescribe equipment and protocols for video teleconferencing on their system, and thereby ensure compatibility and interoperability. However, there is a variety of equipment in use, and establishing a video teleconference with an off-network site may be difficult or impossible due to incompatibility. For instance, there are a variety of protocols for compressed digital video signals, such as proprietary protocols of codec manufacturers CLI and Picturetel. Joining a site having, for example, a Picturetel codec in a video teleconference with a hub and other sites having CLI codecs has been effected by using a Picturetel codec to convert the compressed digital video received from the Picturetel site into analog video, and using a CLI codec to convert the analog video into compressed digital video in the CLI format so that it can be processed by the hub and other sites. This is a cumbersome procedure and requires substantial effort to assemble and configure equipment to conduct a particular video teleconference. As with compressed digital video equipment and protocols, there are a variety of formats for graphics and data which may be required in collaborative computing applications of multimedia teleconferencing. There are also a variety of communication channels and signal formats via which video teleconferencing signals may need to be interchanged. Conference participants at different sites may not want or be able to handle all information generated at every site; for example, a person having only a cellular phone may wish to speak with the other participants in a

collaborative computing video teleconference. The permutations and combinations of these site and equipment variables make it difficult or impossible to conduct many video teleconferences which might be desired.

There are also human variables that affect the interchange of information in a video teleconference. A principal factor is language. A participant may obtain little benefit from a conference in which he does not understand the languages spoken and written by others.

For the foregoing reasons, there is a need for improved systems for multimedia teleconferencing.

SUMMARY OF THE INVENTION

It is therefore a general object of the invention to provide a system which can provide multimedia teleconferencing among a wide variety of sites having a wide variety of equipment types accessible over a wide variety of communication channel types.

In accordance with the invention, the system includes a hub having a plurality of input/output ("I/O") ports, each of which may be coupled to a communication channel for interchanging teleconferencing signals with remote locations. The hub includes a switch that selectively couples I/O ports with each other to set up a teleconference among sites coupled via communication channels to the selected ports, and that may selectively distribute signals to and from the coupled ports. The system further includes selective processing of signals prior to distribution so that the distributed signals are in a form desired by the recipients or required by their equipment. Such selective processing may include video, data, graphics, and communication protocol or format conversion, and language translation. A hub according to the invention may effect such processing by use of signal processors dedicated to specific functions, by programmable signal processors that can perform several functions and are programmed to perform specific functions as needed, or both. A hub according to the invention may generate data relating to the use of its processing capabilities during a teleconference, so that accounting or billing for a

teleconference may be based at least in part on which hub resources were used, the extent of their use, and the person desiring their use. In one aspect of the invention, the identification of a signal processing function to be used during a teleconference is automatically performed in response to the content of signals received at the hub during the teleconference. These and other objects and features of the invention will be understood with reference to the following specification and claims, and the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram illustrating a teleconferencing system including a plurality of sites interconnected by a hub.

Figure 2 is a block diagram illustrating another teleconferencing system including a plurality of sites interconnected by a hub.

Figure 3 is a block diagram illustrating a teleconferencing site.

Figure 4 is a block diagram illustrating a teleconferencing hub in accordance with the invention.

Figure 5 is a block diagram illustrating the data distribution processor of Figure 4 in greater detail.

Figure 6 is a block diagram illustrating the data which may be stored at a teleconferencing site in a system according to the present invention.

Figure 7 is a block diagram illustrating the data which may be stored at a teleconferencing hub in a system according to the present invention.

Figure 8 is a block diagram illustrating the hub controller functions which may be performed by a teleconferencing hub controller in a system according to the present invention.

DETAILED DESCRIPTION

Fig. 1 is a block diagram illustrating an arrangement of elements in a multipoint teleconferencing system. The system includes a plurality of sites, 2A, 2B, 2C, and 2D being illustrated, at which people can participate in a teleconference.

Each of the sites 2 has equipment for receiving electronic signals originating at other sites and generating teleconferencing information outputs for local participants, for generating electronic signals representing locally-generated teleconferencing information for distribution to other sites, or both. Such teleconferencing information may include speech, images, and data, and the electronic signals representing such information may include audio, video, and data signals. Each site 2A, 2B, 2C, and 2D participating in a teleconference is coupled via a communication channel 6A, 6B, 6C, and 6D, respectively, to a hub 4. In order to set up and conduct a teleconference among a set of sites, the hub 4 couples the communication channels from each such site. The hub 4 may also selectively distribute signals to and from the participating sites. Thus a hub may include a voice actuated multipoint control unit, or "MCU", that determines which conference site is generating the dominant audio signal, and switches the video signal from that site to the other sites in the conference. The hub effects a star network topology in which communication channels 6 are "spokes" which carry multimedia teleconferencing signals including digital video signals between the hub 4 and the various sites 2 participating in the conference. While this topology and method of operation are well known prior art, it should be understood that the apparatus and method of the present invention may be used in such a network topology.

A teleconferencing system is relatively easily implemented using the topology of Figure 1 when the sites 2, the communication channels 6, and the hub 4 are under common control. For instance, when the sites 2 belong to a single business entity, that entity can provide the site equipment, provide communication channels suitable for carrying the signals generated by the site equipment, and provide a hub suitable for switching those signals. When all of the variables are under the control of one entity, it can provide whatever capital and whatever operations and administrative effort is necessary to achieve the communications capabilities it desires. By analogy to telephone systems, if the system of Figure 1 were a PBX system used by a business for its internal telephones, it would be relatively easy to provide a desired level of access and interoperability; likewise, prior to deregulation, the U.S. public telephone network

(including terminal equipment at sites 2, lines 6, and switches 4) was under the control of AT&T, which could control access and interoperability. To date, multimedia teleconferencing, particularly multipoint teleconferencing, has generally followed one of the foregoing models. Large companies such as the assignee of the present invention have installed hub and site equipment at their own facilities and have allocated part of their inter-site telecommunications resources to handling teleconferencing signal traffic. Large telecommunications carriers such as AT&T, Sprint, and MCI that already own networks of communication channels and switching hubs have made fixed and portable site equipment available for public use by smaller entities or for conferences among disparate entities.

Figure 2 is a block diagram of a teleconferencing system that is intended to illustrate an environment in which multipoint multimedia teleconferencing may be desired that is more general than that of Figure 1. In the environment of Figure 2, while some sites such as site D may be directly associated with a hub 4, teleconferencing may be desired with other sites A, B, and C having arbitrary types of equipment, communication channel access, and conference requirements. These sites can communicate with hub 4 via cloud 8 via site-associated communication channels 10A, 10B, 10C and hub-associated communication channels 12A, 12B, 12C. Cloud 8 represents the set of non-dedicated communications channels that can be created to interconnect terminal equipment at various sites, and may include portions of the public telephone network, private networks such as local area and wide area networks, and public data networks. The cloud operates to transport signals between a particular site-associated communication channel 10 and a particular port-associated communication channel 12.

Figure 3 is a block diagram illustrating the features which may be present at a teleconferencing site 2 that may participate in a teleconference in accordance with the invention. Teleconferencing-related signals are interchanged with remote locations over site communication channel 10 coupled to the site equipment at site port 72. Site communication channel 10 may comprise an electrical signal channel such as a telephone channel including a POTS line, an ISDN line, a T1 line, or a fractional T1

line, or a network channel such as an ethernet channel; a wireless signal channel such as radio, cellular telephone, or optical or infrared; or a fiber optical channel such as SONET. Signals are interfaced between site port 72 and the site equipment by a port I/O processor 70 appropriate to the communication channel 10, such as a modem, network interface card, or the like.

Site 2 may include a variety of transducers to input teleconferencing information from, and/or output teleconferencing information to, teleconference participants at the site. A multimedia video teleconference may include visually conveyed information, such as video (moving images), graphic information (still images), and alphanumeric or other character-based information, hereinafter referred to as "text". Transducers for such visually conveyed information may include analog transducers such as analog video camera 42 and television 44 that input and output analog video signals 46, and digital video camera 48 and video monitor 50 that input and output digital video signals 52. A multimedia teleconference may include audio information, such as speech. Transducers for such audio information may include analog transducers such as microphone 54 and speaker 56 that input and output analog audio signals 58; A/D converter 60 and D/A converter 62 may be interposed in the signal paths to input and output digital audio signals 64. Digital data representing teleconferencing information may be stored in, input into, and/or output from a memory 66, such as RAM or disk storage in a computer at site 2. This data can be interchanged as formatted digital data 68 with participants at other sites, and can represent stored audio, visual, text, or computer application information. Site signal processor 40 performs signal processing on the signals received from other sites for display at the site 2, and on the signals generated at site 2 for transmission to other sites, the nature of the processing being dependent on the equipment in use at site 2. For instance, site signal processor 40 may include a codec for generating compressed digital video signals from the video signals generated at site 2. Although illustrated as a single block, site signal processor 40 may perform a variety of functions and be implemented using one or several pieces of hardware, the specifics being generally a matter of design choice.

The transducers and site signal processor 40 at a site 2 may provide or require signals representing teleconferencing information in a particular format. For instance, a site in Europe may have analog video equipment 42, 44 operating with analog video signals 46 in the PAL format, while equipment at a site in North America may operate with analog video signals 46 in the NTSC format. The site signal processor 40 may code and decode video in one of a number of compressed digital video formats, such as the open, standards-based MPEG and H.261 formats and the proprietary INDEO, SG3, and Rembrandt formats. Graphics may be in formats such as JPEG, TIFF, GIF, or single frames of a video signal format such as QCIF or CIF. Data may be in a variety of formats such as ASCII or software application - specific types, or may be encrypted. These variations make it difficult to conduct a video teleconference among a randomly selected set of sites having video teleconferencing equipment.

The foregoing factors relate to the technical aspects of acquiring information from teleconference participants, converting the information into electrical signals, preparing the signals for transmission, transmitting, and then the reverse. However, the value of a teleconference is in the effectiveness of its distribution and interchange of information, and the effectiveness of a teleconference in doing so is a strong function of human factors as well as technical factors. For instance, in addition to variations in video, graphics, and like formats from site to site, there are also variations in languages among conference participants. Present day teleconferencing systems do not address these factors.

Figure 4 is a block diagram illustrating a teleconferencing hub in accordance with the invention which facilitates teleconferences among sites with disparate equipment and participants. The hub 4 includes a plurality of hub ports 22 to communicate with teleconference sites. Teleconferencing-related signals are interchanged with remote locations over hub communication channels 12 coupled to the site equipment at hub ports 22. As with the site communication channels 10, a hub communication channel 12 may comprise an electrical signal channel such as a telephone channel including a POTS line, an ISDN line, a T1 line, or a fractional T1 line, or a network channel such as an ethernet channel; a wireless signal channel such

as radio, cellular telephone, or optical or infrared; or a fiber optical channel such as SONET. Signals are interfaced between a hub port 22 and the primary hub signal processing equipment by a port I/O processor 24 appropriate to the communication channel 12, such as a modem, network interface card, or the like.

Teleconference signals received at each port, after processing by the port I/O processors 24, are presented to data distribution processor 26. Data distribution processor 26 performs, among other things, the basic signal routing functions required for a multipoint teleconference. These functions include the selection and distribution of teleconferencing signals received from certain sites to others participating in a teleconference, such as voice actuated selection of video signals. Such functions are described more fully below with respect to Figure 5. Illustrated in Figure 4 is an important aspect of the invention, namely, the processing of teleconferencing signals received from one site to facilitate effective communication of the information contained in those signals to participants at other sites. Such processing can be performed in several ways. Figure 4 shows a plurality of programmable signal processors 28, which may be selectively disposed to process signals received at one port for distribution to one or more other ports. The programmable signal processors 28 process input signals in the manner specified by stored software programs indicated in Figure 4 as signal processing utilities 30. Figure 4 shows several types of signal processing utilities 30 that may desirably be used in a system according to the invention. It should be understood that for each type of utility provided, there will desirably and in general be a plurality of specific utilities provided, each of which may be invoked by loading it into a programmable signal processor 28 to perform a specific signal processing function of the general type indicated.

Thus, codec utility 30A provides the video codec functions of converting analog audio and video signals into compressed digital video signals, and converting compressed digital video signals into analog audio and video signals. Specific codec utilities 30A would be provided in a system to account for the variety of analog audio and video signal formats and compressed digital video signal formats that are to be handled. Thus, for example, there may be NTSC to H.261, NTSC to INDEO, PAL to

H.261, and PAL to INDEO codec utilities, and as many permutations and combinations of formats as are desired.

When the codec function is performed in a programmable signal processor 28 that is running a codec utility 30A, the hub is functioning with a software codec. At present, in video teleconferencing systems, whether to use a software codec or a hardware codec is a matter of design choice, determined by the availability, cost, and effectiveness of software and hardware codecs to perform a specific function, and these factors change as the field evolves. The same is true with the system of the present invention. Thus, Figure 4 includes dedicated signal processors 32 for those functions that are better performed in hardware. Thus, for example, a system might use a software codec comprising a programmable signal processor 28 and a codec utility 30A as a PAL to INDEO codec, and a hardware codec comprising a dedicated signal processor 32 as a NTSC to H.261 codec. The system of the present invention is thus flexible in that as research, development, and product introductions proceed, as standards change, and as the availability, cost, and effectiveness of software and hardware to perform the various signal processing functions change, the hub may be reconfigured to provide desired functions and optimize their implementation.

Transcoding refers to direct conversion of signals from one compressed digital video format to another. The signal processing utilities 30 of Figure 4 include a transcoding utility 30B to perform this function. Specific transcoding utilities 30B may be provided, for example, for transcoding INDEO to H.261, SG3 to Rembrandt, and any other desired now-existing or later developed compressed digital video format. As with the codec function described above, and as with all the signal processing functions indicated in block 30, the transcoding function may be performed in a programmable signal processor 28 or in a dedicated signal processor 32, as desired.

Coding recognition utility 30C may be provided to enable the hub 4 to automatically determine the format of a signal received at a hub port 22. If the received signal is intelligible as a signal in a format the hub 4 can process, the hub 4

can automatically invoke the signal processing resources needed to account for differences in signal formats in a particular teleconference.

Speech translation utility 30D provides for translation of audio speech from one natural language to another. Thus, for example, speech translation utilities 30D may be provided to translate English speech to German, Russian speech to French, and the like. This enables, for instance, participants at a site in the United States to speak English and have the audio output to participants at a site in Germany rendered with a German speech translation of the English, and vice versa. The speech translation and codec functions that have been discussed illustrate an important aspect of the present invention. That is, while many functions can be implemented either at a site or at a hub, some are far more advantageously implemented at a hub. It may be feasible, for instance, to have desktop systems at a site which are capable of performing several types of codec or transcoding functions. However, even if it were presently feasible and cost-effective to do so, as compressed digital video formats proliferate and evolve, it would likely be impractical to continually upgrade site equipment to maintain its capability to teleconference with the then-existing equipment at other sites. Since multipoint teleconferences require a hub or bridge function in any event, it is preferable to provide the relatively few hubs with high-powered and expensive functionality, and keep them current with the state of the art, than to attempt to do so with the relatively many site-based systems. This is also the case with speech translation functions. It may be feasible at present or in the near future to provide a desktop or site-based server system with the ability to perform sufficiently real-time translation of speech from one or several natural languages to one or several others. The capabilities of computer hardware continue to increase exponentially, and the field of natural language processing is active as well, and so it may become possible in the future to provide site systems with the capability to translate between all human languages and dialects. It is believed that those capabilities will be achieved, if at all, only in the far future. However, even if it were possible to implement them, it would be undesirable to do so, since there would be many sets of languages that a particular site would seldom, if ever, be called upon to

handle in a teleconference. An advantage of the present invention is that while the signal processing capabilities may be required or economically justifiable for a hub, once provided they may be highly useful in other circumstances. For instance, two sites with incompatible video codecs may be able to hold a point-to-point video teleconference by routing their call through the hub of the present invention. Or two persons who do not share a common language can understand each other in a point-to-point telephone call by routing it through the hub. Two persons who do not share a common language and who are face-to face with each other can even hold a conversation and understand each other by use of the present invention: rather than talk directly to each other, they could place a telephone call through the hub and invoke the appropriate natural language speech translation function. Thus, for example, if a hub according to the present invention were accessible through the U.S. public cellular telephone network, a foreign language speaking tourist could carry a cellular phone, enter a U.S. shop and discuss an item of merchandise with a proprietor who speaks only English by a telephone call routed between the tourist's cellular phone and the proprietor's phone via the hub.

Although it may be economically justifiable in principle or in general to provide a hub with the signal processing functions illustrated in Figure 4, the justification in a particular instance may depend on the ability of a hub operator to recover costs. As shall be discussed more fully below, a function may be more easily justified and its costs recovered if the hub operator can perform usage measurement and usage-based billing for the functions provided. It may also be noted at this point that although the block diagram of Figure 3 may suggest that the hub 4 shown therein is a single item located in a single place, this is not necessarily the case. The hub function of Figure 3 may be provided by several hubs of the sort illustrated in Figure 4, and these hubs may be collocated or at different facilities, and operated by the same or different hub operators. The hubs may be directly coupled by dedicated communication channels, or may be coupled as needed for particular teleconferences, such as via hub ports 22. For example, a teleconference might be mediated by a first hub which lacks the ability to translate between particular languages desired by the

participants. The first hub can couple to a second hub which does have the desired translation capability, and selected signals communicated to the second hub for processing. If the two hubs are operated by different entities, means may be provided for inter-company settlement of accounts, including means based on usage measurement and usage-based billing. The invocation of the signal processing resources of one hub by another may be made to occur automatically, such as by querying other candidate hubs for availability and establishing a communication channel for teleconferencing signals with the first available hub. The selection process may be qualified, such as in the event that a conference participant has a preference for a particular signal processor's accuracy of speech translation or realism of codec function, or to select the lowest cost of conducting the desired teleconference.

Continuing with the description of signal processing utilities 30 that may be provided in accordance with the present invention, Figure 4 shows text translation utility 30E that provides translation between text information rendered in different languages. This may involve translation between words written in different natural languages, for instance between English writing and German writing. Another type of such text translation is between computer file types, such between ASCII and EBCDIC characters, which may be useful in teleconferences involving collaborative computing.

Figure 4 also shows speaker recognition utility 30F. This utility analyzes the audio signals received during the course of a teleconference to select speech signals and to identify who the speaker is. This function can have several uses in a teleconference. If the hub contains stored data representing the speech characteristics of a participant, then the speech translation function can synthesize the audio representing the translated speech so that the tone of "voice" sounds like the particular speaker is actually speaking in the foreign language. If speech translation is used, the selection of which teleconferencing signals are to be directed to which signal processor can be easy. For instance, referring to Figure 2, if all participants at site A speak English and all participants at site B speak German, the hub 4 simply routes all

audio speech signals received over communication channel 12A to an English to German signal processor, and routes all audio speech signals received over communication channel 12B to a German to English signal processor. However, if a Japanese-speaking person is also a participant at site A, this simple method cannot work. To effectively conduct the teleconference, the hub 4 should not only deliver speech audio signals to site A translated into both English and Japanese, it must also translate both English and Japanese from site A into German. One way to accomplish this is to provide entirely separate audio circuits at site A, one for English speech and one for Japanese speech, and supply separable audio signals to the hub. This may be difficult in some cases, in which event the selection of the translator to which the speech signal should be routed may be done based on analysis of the contents of the speech signal itself. It is doubtful that an identification of language from the speech per se could be performed quickly enough to do acceptably near-real-time translation. However, identification of a speaker from the voice characteristics of speech can be done relatively quickly, particularly from a limited set of candidate speakers in a particular teleconference. Thus, the hub 4 of the present invention can quickly determine, using speaker recognition utility 30F, that the person who just started speaking is Mr. Jones, and relying on stored data that identifies Mr. Jones as English speaking, the hub can direct speech-containing audio signals then being received from site A to the English-to-German speech translation utility 30D. Another use of the speaker identity yielded by speaker recognition utility 30F is in what might be termed "intelligent" audio-based switching of the video signals. At present, MCU's determine the received compressed digital video signal having the "loudest" audio component, and select that signal to be distributed to the other sites so that the participants there can see the person who is speaking at any given time. This method makes the assumption that whoever speaks loudest deserves to be heard, which is sometimes but not always the case. It may be desired to have only a subset of the participants who are to be focused on, and a priority of importance among the members of the subset. For instance , if the company president and his administrative assistant are at Site A, it may be desired to distribute the video from site A whenever the president is speaking,

no matter how quietly, and not distribute the video (and perhaps audio) from site A when the administrative assistant is speaking, no matter how loudly. By identifying the president and the administrative assistant using the speaker recognition utility 30F, the system of the present invention can effect such speaker-dependent video switching.

To identify a speaker based on the voice characteristics of speech, the hub 4 must maintain stored data representing the voice characteristics of the conference participants. If the hub does not contain such data for a particular participant, it can be obtained during initiation of the teleconference. For example, when site A connects to the hub, a new participant might say "I'm John Doe, president of XYZ Corporation, and I speak English", and data regarding his language preferences and voice characteristics can be derived from this speech and stored at the hub.

Figure 5 is a block diagram illustrating the data distribution processor of Figure 4 in greater detail. A controller 38 provides control signals 112 to the port I/O processors 24 to set up communications over communication channels 12 to be used in a teleconference. The teleconferencing signals 100 received at hub ports from teleconference sites are supplied to a signal processor switch 34. Controller 38 provides control signals 110 including signal processor switch control signals in order to select signal paths for the received signals. If no processing is required for particular received teleconferencing signals, the unprocessed teleconferencing signals 106 are routed directly to signal distribution switch 36 as inputs to it. If processing is required for particular received teleconferencing signals, they are routed by signal processor switch 34 to appropriate signal processors as input signals 102 thereto. Although Figure 5 shows the available signal processors as three programmable signal processors 28, it should be understood that this is merely for purposes of illustration and a system in accordance with the present invention may provided with as many programmable signal processors 28 and as many dedicated signal processors 32 as are required to perform the desired signal processing functions. If the signal processing requires use of a programmable signal processor 28, control signals 114 cause a selected programmable signal processor 28 to load and launch the appropriate signal

processing utility 30. The processed output signals 104 of the selected signal processors are supplied as inputs to signal distribution switch 36. In response to control signals 110 received from controller 38, signal distribution switch 36 distributes selected of its input signals 104, 106 as distribution switch output signals 108 to selected hub ports, where they are communicated to remote sites. The controller 38 includes a processor 80 which executes controller applications 84 and reads data from and writes data to a memory 82.

Figure 6 is a block diagram illustrating the data which may be stored in a site memory 66 at a teleconferencing site in a system according to the present invention. Such data may be generally categorized as site data 90 relating to the site and its equipment, participant data 92 relating to the people who have or may participate in a teleconference at the site, and reservation and scheduling data 94. Site data 90 includes site identifying data 90A; site equipment identifying data 90B that can be transmitted to the hub during conference reservation or setup to enable the hub to invoke the appropriate codec, transcoding, and text translating signal processors; site billing information 90C identifying who should be billed for teleconferences conducted from the site; and a site address file 90D containing the identities and communication channel addresses of a list of other sites that have or may be expected to participate in teleconferences with the site. Participant data 92 includes records for each person who has been or may be expected to be a participant in a teleconference at the site. Each participant record may include fields containing data representing the participant's name (92A) , voice characteristics (92B), languages spoken (92C), identities and communication channel addresses of other sites that have or may be expected to participate in teleconferences with the participant (92D), and participant billing information (92E) identifying who should be billed for teleconference in which the person is a participant. The memory 66 may contain reservation and scheduling data 94 representing the time a teleconference is to take place, the sites at which it will take place, and the persons who will be participants. If the foregoing data is stored in the site memories 66 of the teleconferencing sites which are to participate in a teleconference, then the sites can communicate that data to the hub during reservation

and scheduling of a teleconference or during initiation of the teleconference. Thus, for example, a site which desires to schedule a teleconference and reserve hub resources for the teleconference can do so by contacting the hub, transmitting its reservation and scheduling data 94, transmitting its site identifying data 90A, transmitting its site equipment identifying data 90B to enable the hub to reserve signal processing resources required for the site equipment, transmitting its site billing information 90C to enable the hub to properly bill for the planned teleconference, transmitting from its site address file 90D selected addresses for the other sites identified in the reservation and scheduling data 94 as sites that will participate in the teleconference to enable the hub to call them, and transmitting participant data 92 for the persons identified in the reservation and scheduling data 94 as participants, to enable the hub to reserve signal processing resources.

Figure 7 is a block diagram illustrating the data which may be stored at a teleconferencing hub in a system according to the present invention, and Figure 8 is a block diagram illustrating the hub controller functions which may be performed by a teleconferencing hub controller in a system according to the present invention. Data stored in hub memory 82 may include site and participant data 82A, 82B, 82C pertaining to sites A, B, and C, respectively. Hub 4 may acquire this data by periodically mirroring the site memory 66 of each site which has participated or is expected to participate in a conference mediated by the hub, or by creating or updating records contained in its memory 82 every time the hub receives data from a site during scheduling or conduct of a teleconference. To conduct a teleconference with persons or sites that have not previously participated in a teleconference mediated by the hub, the hub acquires and stores new site/participant data 82D during scheduling or setup of the teleconference. Data stored in hub memory 82 may include terminal and communication channel related data 82E, which is used by terminal recognition application 84B and channel recognition application 84C. Such data may include for instance telephone numbers, which if received by caller-ID functionality in the hub can be quickly associated with site and participant identities; identification of dedicated communication channels likewise can be quickly associated with site and

participant identities. A usage measurement application 84G generates and stores usage data 82F by monitoring which of its resources (such as signal processors) are used for how long for which site or participant. A billing application 84H generates and stores billing data 82G based upon usage data 82F and stored data representing predetermined billing parameters, the billing data 82G representing who should be billed how much for the use of hub resources. The billing parameters may be selected to reflect the relative costs of the hub resources. For example it may be extremely expensive to provide hardware and software to perform Swahili to Sioux speech translation, and the costs would have to be recovered from very infrequent invocation of this function. To recover the cost, the billing parameters would reflect a very high price per unit time of use.

As has been noted, a hub may lack certain resources required to conduct a requested teleconference, and remote resource data 82I stored in memory 82 may be used to locate, contact, and schedule resources located in other hubs or sites. In that event, the usage measurement application 84G and billing application 84H may generate usage data and billing data representing the use of these remote resources

Controller applications 84 include encryption/decryption application 84D which may be used if sites 2 are transmitting encrypted teleconferencing signals. This will most likely be desired with text information, but may be done with other or all components of a teleconferencing signal. Encryption/decryption application 84D decrypts the site signals so they can be processed by hub 4, and encrypts them for transmission back to the sites. If different participating sites use different encryption/decryption methods, encryption/decryption application 84D may be required to translate between them, in a manner analogous to that required when sites have different video codecs. Although the encryption/decryption function could be performed by a programmable signal processor 28 or a dedicated signal processor 32, it may be necessary to decrypt a received site signal before any other processing is done, and so Figure 5 and Figure 8 illustrate this function as being performed by the controller 38.

In view of the foregoing, it is seen that the system of the present invention provides great flexibility in mediating teleconferences, particularly multimedia multipoint teleconferences, among disparate sites having incompatible equipment and participants lacking common language. The system of the present invention achieves these results by providing a plurality of signal processing functions and selecting the teleconferencing signals to be processed and the signal processing functions to be applied so as to bridge the barriers posed by site incompatibility. The selection of the teleconferencing signals to be processed and the signal processing functions to be applied is made by the hub at least in part in response to signals received from the participating sites; the received signals to which the hub is responsive may be "handshaking" signals interchanged with the hub to initiate a teleconference, signals interchanged among the sites during the course of the teleconference, or both. The system of the present invention delivers teleconferencing signals to a site that are optimized in form and in content to convey information to the particular participants at the particular site.

Variations on the systems disclosed herein and implementation of specific systems may no doubt be done by those skilled in the art without departing from the spirit and scope of the present invention.

WHAT IS CLAIMED IS:

1. A hub system for teleconferencing comprising:

   a plurality of ports each adapted to be coupled to a communication channel to receive teleconferencing signals therefrom and transmit teleconferencing signals thereto;

   a plurality of signal processors, each adapted to receive teleconferencing signal inputs and process said teleconferencing signal inputs to provide processed teleconferencing signal outputs;

   a processor switch coupled to said ports and to said signal processors for selectively coupling teleconferencing signal inputs to said signal processors in accordance with control signals;

   a distribution switch coupled to said ports and to said signal processors for selectively coupling processed teleconferencing signal outputs from said signal processors to said ports in accordance with control signals; and

   a controller coupled to said switches for providing control signals thereto, thereby selectively controlling the processing and distribution of teleconferencing signals received at said hub.

2. A system according to claim 1, wherein said signal processors include video codecs.

3. A system according to claim 1, wherein said signal processors include video transcoders.

4. A system according to claim 1, wherein said signal processors include natural language speech translators.

5. A system according to claim 1, wherein said signal processors include text translators.

6. A system according to claim 1, wherein said controller generates said control signals in response to teleconferencing signals received at said hub from said sites..

7. A system according to claim 1, wherein said hub analyzes a speaker's voice characteristics to identify the speaker, and said signal processors process said teleconferencing signals in response the identity of the speaker.

8. A system according to claim 1, wherein said hub measures the usage of said signal processors and generates billing data based on said measured usage.

9. A method for conducting a teleconference among three or more conference sites comprising the steps of:

providing a hub, said hub including a plurality of available signal processing functions that may be selectively performed on teleconferencing signals received at said hub from said sites;

selecting a set of received teleconferencing signals to be processed and a set of said available signal processing functions to be performed on said received teleconferencing signals to be processed;

performing said selected signal processing functions on said selected received teleconferencing signals; and

transmitting said processed teleconferencing signals to one or more of said sites.

10. The method of claim 9, wherein said signal processing functions include a video codec function.

11. The method of claim 9, wherein said signal processing functions include video transcoding.

12. The method of claim 9, wherein said signal processing functions include natural language speech translation.

13. The method of claim 9, wherein said signal processing functions include text translation.

14. The method of claim 9, wherein said signal processing functions include audio speech analysis to identify a speaker.

15. The method of claim 9, further including the step of monitoring said performing step to generate usage data reflecting the extent of use of said selected signal processing functions.

16. A system for teleconferencing comprising a hub and a plurality of remote sites each having site teleconferencing equipment coupled to said hub by a communication channel, wherein said hub includes:

a plurality of ports each adapted to be coupled to a communication channel to receive teleconferencing signals therefrom and transmit teleconferencing signals thereto;

a plurality of signal processors, each adapted to receive teleconferencing signal inputs and process said teleconferencing signal inputs to provide processed teleconferencing signal outputs;

a processor switch coupled to said ports and to said signal processors for selectively coupling teleconferencing signal inputs to said signal processors in accordance with control signals; and

a distribution switch coupled to said ports and to said signal processors for selectively coupling processed teleconferencing signal outputs from said signal processors to said ports in accordance with control signals,

wherein said processor switch is responsive at least in part to control signals derived from signals received by said hub at one or more of said ports from one or more of said sites, whereby said sites can selectively control the processing of teleconferencing signals at said hub.

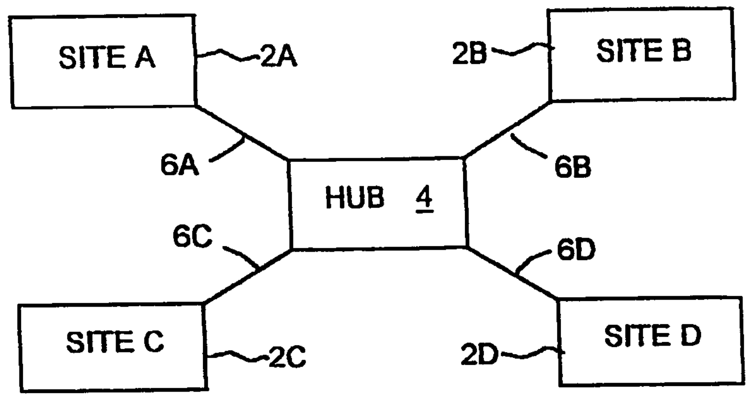
SITE A
2A
2B
SITE B
6A
HUB 4
6B
6C
6D
SITE C
2C
2D
SITE D

Fig. 1

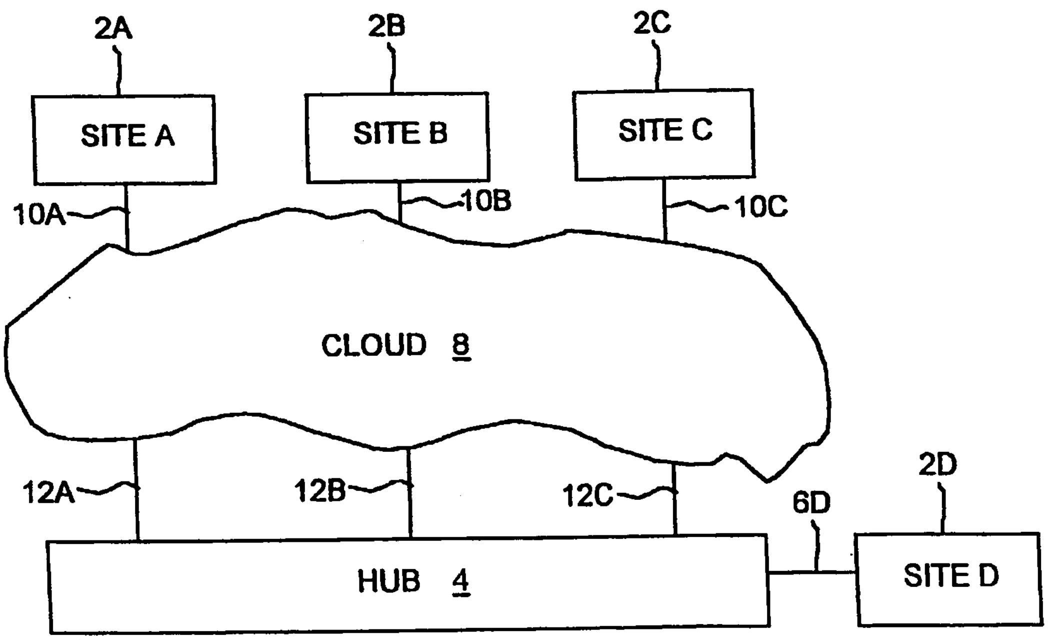
2A
2B
2C
SITE A
SITE B
SITE C
10A
10B
10C
CLOUD 8
12A
12B
12C
6D
2D
HUB 4
SITE D

Fig. 2

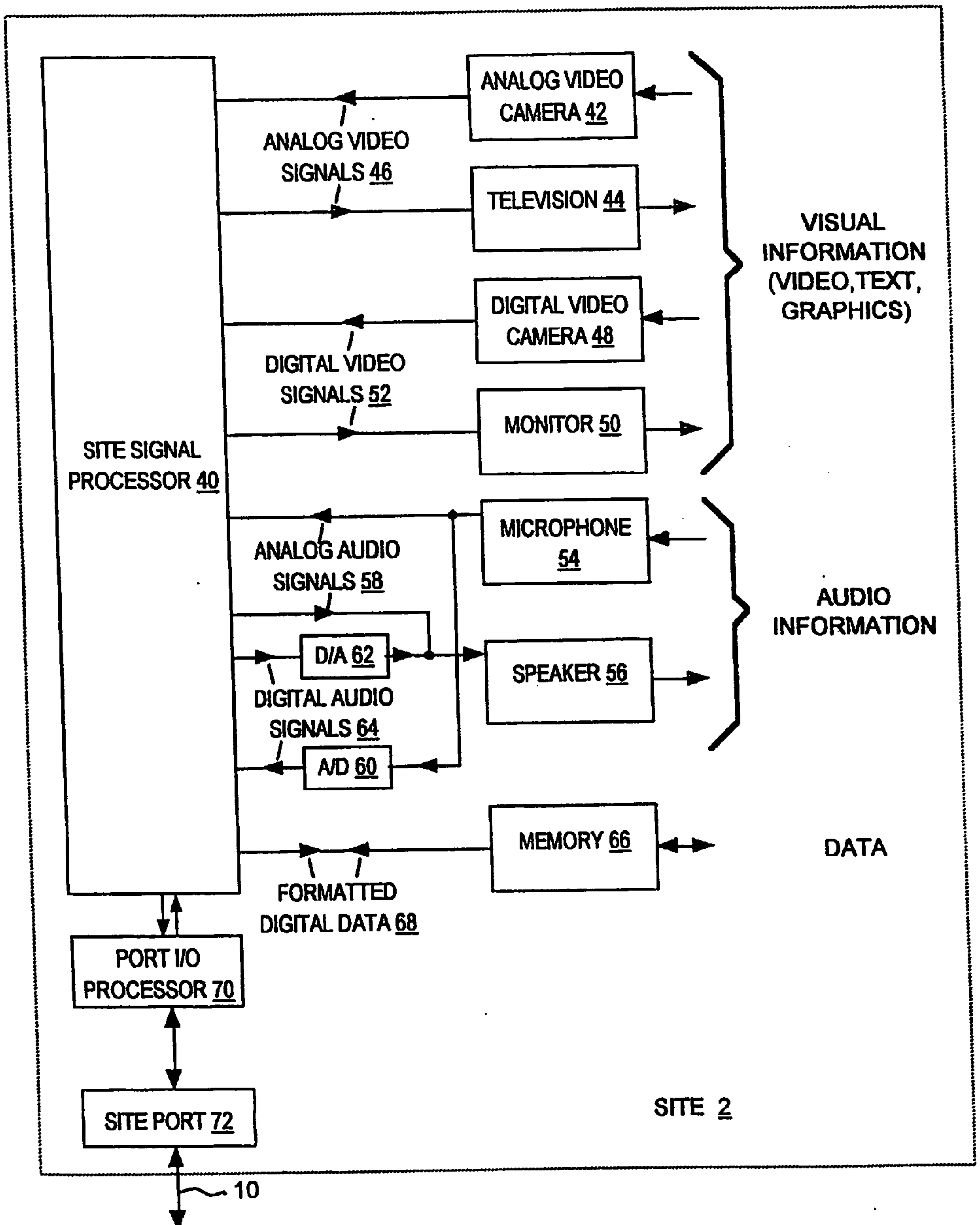
SITE SIGNAL PROCESSOR 40
ANALOG VIDEO CAMERA 42
ANALOG VIDEO SIGNALS 46
TELEVISION 44
VISUAL INFORMATION (VIDEO,TEXT, GRAPHICS)
DIGITAL VIDEO CAMERA 48
DIGITAL VIDEO SIGNALS 52
MONITOR 50
MICROPHONE 54
ANALOG AUDIO SIGNALS 58
AUDIO INFORMATION
D/A 62
SPEAKER 56
DIGITAL AUDIO SIGNALS 64
A/D 60
MEMORY 66
DATA
FORMATTED DIGITAL DATA 68
PORT I/O PROCESSOR 70
SITE PORT 72
SITE 2
10

Fig. 3

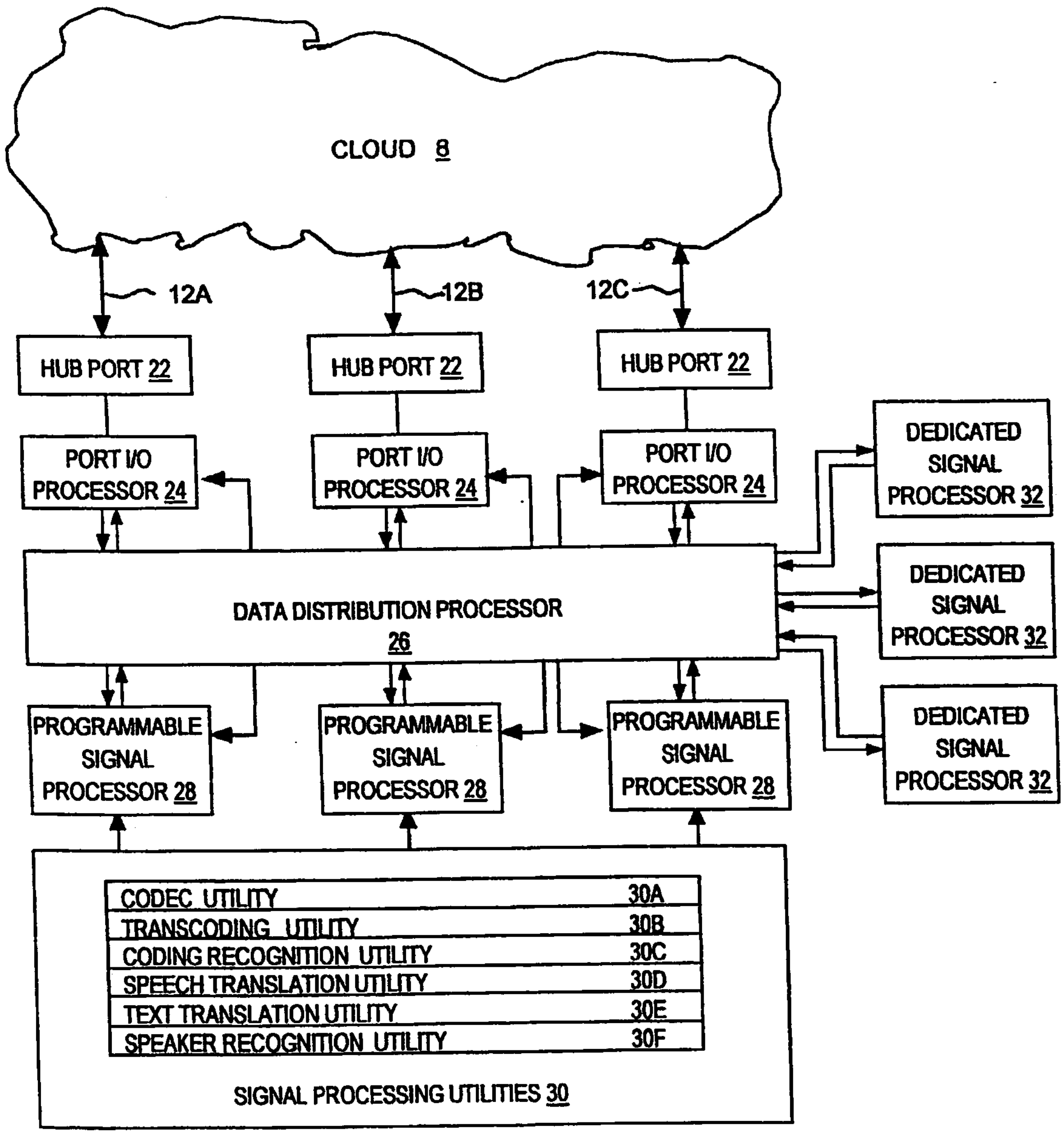
CLOUD 8
12A
12B
12C
HUB PORT 22
HUB PORT 22
HUB PORT 22
PORT I/O PROCESSOR 24
PORT I/O PROCESSOR 24
PORT I/O PROCESSOR 24
DEDICATED SIGNAL PROCESSOR 32
DEDICATED SIGNAL PROCESSOR 32
DEDICATED SIGNAL PROCESSOR 32
DATA DISTRIBUTION PROCESSOR 26
PROGRAMMABLE SIGNAL PROCESSOR 28
PROGRAMMABLE SIGNAL PROCESSOR 28
PROGRAMMABLE SIGNAL PROCESSOR 28
CODEC UTILITY 30A
TRANSCODING UTILITY 30B
CODING RECOGNITION UTILITY 30C
SPEECH TRANSLATION UTILITY 30D
TEXT TRANSLATION UTILITY 30E
SPEAKER RECOGNITION UTILITY 30F
SIGNAL PROCESSING UTILITIES 30

Fig. 4

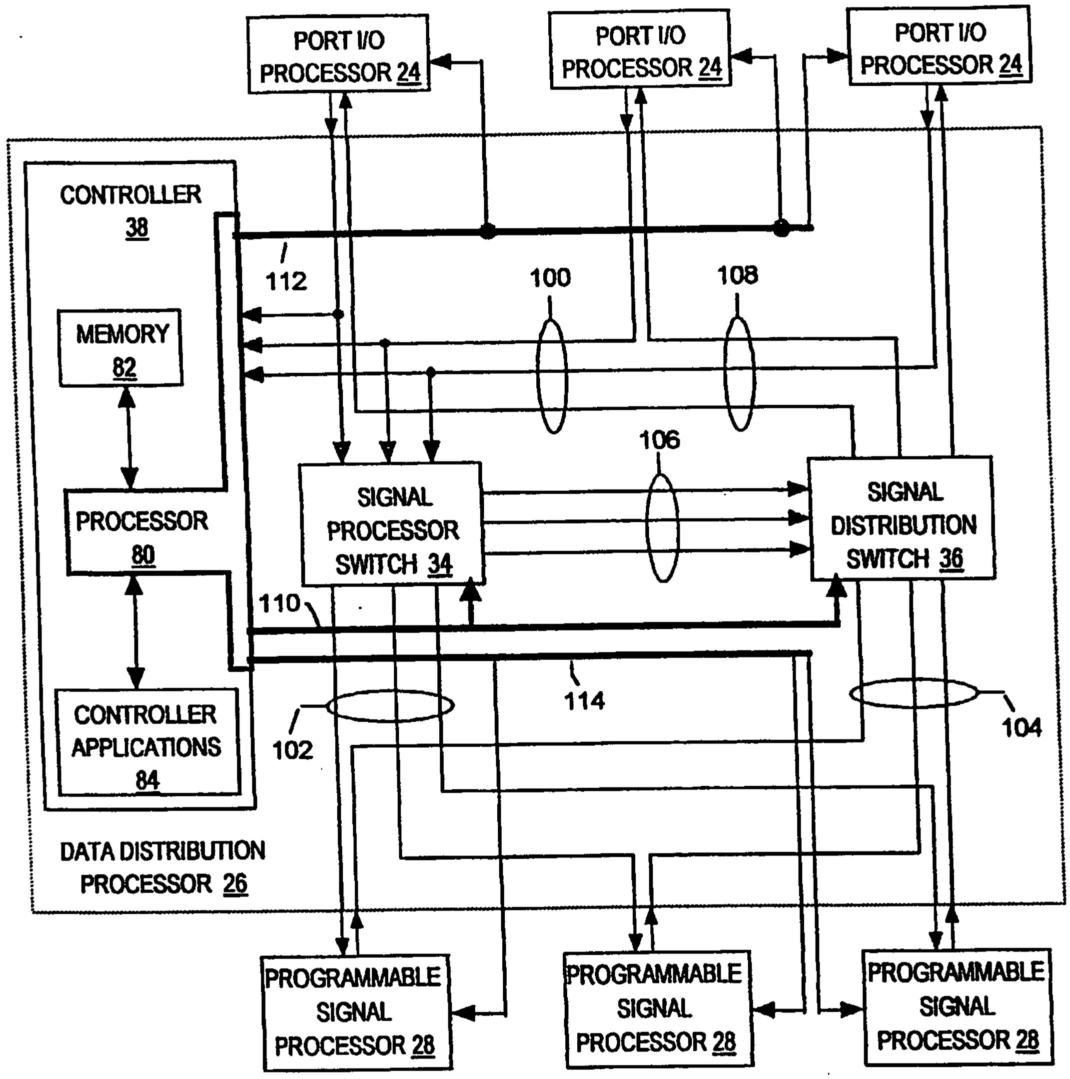
PORT I/O PROCESSOR 24
PORT I/O PROCESSOR 24
PORT I/O PROCESSOR 24
CONTROLLER 38
MEMORY 82
PROCESSOR 80
CONTROLLER APPLICATIONS 84
DATA DISTRIBUTION PROCESSOR 26
112
100
108
106
110
114
102
104
SIGNAL PROCESSOR SWITCH 34
SIGNAL DISTRIBUTION SWITCH 36
PROGRAMMABLE SIGNAL PROCESSOR 28
PROGRAMMABLE SIGNAL PROCESSOR 28
PROGRAMMABLE SIGNAL PROCESSOR 28

Fig. 5

SITE ID	90A
SITE EQUIPMENT ID	90B
SITE BILLING INFO	90C
SITE ADDRESS FILE	90D

— SITE DATA 90

NAME	92A
VOICE CHARACTERISTICS	92B
LANGUAGES	92C
ADDRESS FILE	92D
BILLING INFO	92E

— PARTICIPANT DATA 92

RESERVATION/SCHEDULING DATA 94

SITE MEMORY 66

Fig. 6

SITE A SITE/PARTICIPANT DATA	82A
SITE B SITE/PARTICIPANT DATA	82B
SITE C SITE/PARTICIPANT DATA	82C
NEW SITE/PARTICIPANT DATA	82D
TERMINAL & CHANNEL DATA	82E
USAGE DATA	82F
BILLING DATA	82G
RESERVATION/SCHEDULING DATA	82H
REMOTE RESOURCE DATA	82I

HUB MEMORY 82

Fig. 7

COMMUNICATIONS/HANDSHAKING	84A
TERMINAL RECOGNITION	84B
CHANNEL RECOGNITION	84C
ENCRYPTION/DECRYPTION	84D
PROCESSOR SWITCH CONTROL	84E
DISTRIBUTION SWITCH CONTROL	84F
USAGE MEASUREMENT	84G
BILLING	84H
RESERVATION/SCHEDULING	84I

CONTROLLER APPLICATIONS 84

Fig. 8